



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Temporal structure in emerging language

**Citation for published version:**

Schouwstra, M 2017, 'Temporal structure in emerging language: From natural data to silent gesture', *Cognitive Science: A Multidisciplinary Journal*, vol. 41, no. S4, pp. 928-940.  
<https://doi.org/10.1111/cogs.12441>

**Digital Object Identifier (DOI):**

[10.1111/cogs.12441](https://doi.org/10.1111/cogs.12441)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Cognitive Science: A Multidisciplinary Journal

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



*(Paper accepted for publication in Cognitive Science)*

# **Temporal structure in emerging language: from natural data to silent gesture**

**Marieke Schouwstra**

**Centre for Language Evolution, School of Philosophy, Psychology & Language Sciences,  
University of Edinburgh,**

**3 Charles St, Edinburgh EH8 9AD**

**United Kingdom**

**[Marieke.Schouwstra@ed.ac.uk](mailto:Marieke.Schouwstra@ed.ac.uk)**

**Word count: 4100**

**Corresponding author:**

**Marieke Schouwstra**

**[Marieke.Schouwstra@ed.ac.uk](mailto:Marieke.Schouwstra@ed.ac.uk)**

**+44 753 4319811**

**Key words:**

**silent gesture; semantics; temporality; language evolution; word order; tense**

## 1. The emergence of temporal structure in human language

All languages in the world have means to talk about the past and the future. In many languages, this is done by inflection on the verb, and generally, quite some morphological complexity is reserved for temporal structure (see e.g. Dahl & Velupillai, 2013; Dahl, 1985). Moreover, humans have been characterised as being unique (compared to other species) in spending time thinking about things that are not here and now (Hurford, 2007; Millikan, 2004, Gärdenfors 2004). Even though past and future play an important role in human thinking and human communication, little is known about how languages got such sophisticated means for talking about temporally remote events and what the cognitive basis of these linguistic means could be. One way of answering this question is by looking at what people do in the absence of language. In the literature, different instances of such situations are described, such as the language of unsupervised adult second language learners (Klein & Perdue, 1997; Klein, 2001), homesign (Goldin-Meadow & Mylander, 1998; Goldin-Meadow, 2003) and emerging sign languages (Padden, Meir, Sandler & Aronoff, 2010; Senghas, Kita & Özyürek, 2004). The linguistic systems that emerge in these situations are structurally simpler than full languages but they are used successfully for communication. The kind of evidence that is thus obtained may help us understand which properties are fundamental to human communication (Goldin-Meadow, 2003). This paper will investigate emerging language systems to get an idea of *whether* temporal information can be conveyed with modest structural means, and *how* this can be done.

I will start by looking at the semantic and syntactic properties of temporal utterances, then look at observations from natural data, showing that people do talk about past and future, even when they have restricted morphosyntactic means, in a way that mirrors the semantic structure of such utterances. Subsequently, I will show evidence from silent gesture, a novel laboratory paradigm in which hearing people are asked to convey information while using only their hands. By showing that silent gesture triggers similar strategies for structuring

information as those observed in homesign and unsupervised second language learning, I bridge a gap between different strands of research, to obtain a more complete picture of the foundations of temporal structure in human language.

### 1.1 Temporal information in modern languages: semantically separate, but syntactically interwoven

What does it mean to talk about the past or the future? What kind of information is being transmitted? A sentence like ‘Mary will repair her bicycle’ shares certain properties with a sentence like ‘Mary has repaired her bicycle’: they both address an event in which Mary repairs her bicycle, but they differ from each other where it comes to the time of the event. Semanticists describe the former as propositional information, and the second as temporal information. Simple propositions can be seen as a semantic whole: a predicate plus its arguments (the action, *repairing*, plus its participants, *Mary* and *her bicycle*, in the case of the examples above). Temporal information can be divided into information about *tense* and *aspect*. Tense information specifies *when* an action occurs, and can thus be seen as information added to (tenseless) simple propositions. Aspectual information tells us something about the *internal constituency of actions, events, states, processes or situations* (see, e.g., Hamm & Bott, 2014 or Steedman, 1994), but will not be the focus of this paper.

Existing languages have various means to describe tense and aspect, but many languages express it as inflection on the verb (Bybee, Perkins, & Pagliuca, 1994): structurally, temporal information is deeply interwoven with propositional information, and does not reflect the basic semantic organisation of the information. This paper will show that emerging language systems can encode temporal information, but they do this differently from full, existing languages: temporal information is encoded in a way that directly reflects its semantic structure: propositional information and temporal information are represented separately, and

do not influence each other. The temporal information takes scope over the propositional information, and comes first in the utterance.

## 1.2 Evolutionary insights from emerging language

In the process of acquiring a second language outside the classroom, adult learners go through a stage called Basic Variety, which is characterised as follows. It is determined by a small number of organisational principles, largely independent of source and target language,<sup>1</sup> and simple, but efficient for communication (Klein & Perdue, 1997). Utterances in the Basic Variety typically consist of a predicate plus its arguments, but verbs are usually not inflected, or at least not in an informed way,<sup>2</sup> so the standard means for expressing temporal structure are not available. The Basic Variety *does* have a relatively large repertoire of temporal adverbials, and these are recruited when speakers refer to events in the past or future: they are placed in front of an utterance.

Some examples:

(1)

gisteren ik bergen gaan naar

*yesterday I mountains go to*

*(Yesterday, I went to the mountains)*

(Starren 2001, p. 149)

(2)

(interviewer: what are you doing here? Are you working?)

avant je [travaj] / maintenant non

---

<sup>1</sup> Klein and Perdue (1997) do note an influence, for example of source language basic word order (see footnote 6 in their paper), but they stress that this is rare overall.

<sup>2</sup> E.g., learners of English mostly use verb stems (or occasionally a ‘V-ing’ form), and learners of French and German use the infinitive. Speakers may use inflected forms occasionally, but do not do so structurally.

*before I work / now no*

(Benazzo 2009 p. 29)

This shows that even though there is no informed usage of inflection on the verb, temporal information is indeed expressed, by employing adverbials or noun phrases at a sentence initial position. The function of the adverbial is to create a temporal frame: a place on the timeline in the context of which the rest of the information is interpreted. Adverbs most suitably create a temporal frame when they are placed utterance-initial (de Swart, 1999).

Similar structuring is found in an emerging language form in the manual domain: homesign. Homesign systems emerge when deaf children are born into hearing families, and the hearing family members do not know any conventional sign languages. Deaf children will start to spontaneously develop a gestural communicative system. The properties of various homesign systems have been observed and compared (Goldin-Meadow, 2003; Goldin-Meadow et al. 2009) and it was noted that homesign systems developed in different cultures show structural similarities (e.g., a preference to place Patients before Actions, which is consistent with OV sentence structure).

Homesigning children do refer to the non-present (although they start doing it later than hearing children), first by referring to non-present objects and events, and eventually to events in the remote past and future, as well as hypothetical and fantasy events (Morford and Goldin-Meadow, 1997). Because children are still developing cognitively, a better comparison for the Basic Variety data discussed above can be found in adult homesign. Data from adult homesigners is relatively sparse, but an overview is given in Benazzo (2009), who discusses data from Fusellier (2004) from 3 deaf homesigners living in Brazil. One of their ways to refer to events in past and future is, quite similarly to the BV, by using adverbials sentence-initially, such as in the following two examples.

(3)

(about working situation)

before [pointing over shoulder] GOOD

now [pointing downwards] SO SO (more or less)

(4)

(about a football team)

before [rotating hand backwards] WIN [sign for win/victory]

now [pointing downwards] FINISH

(Fusellier, 2004: 282, as quoted in Benazzo, 2009)

Conversely, the temporal expression seems to be utterance final when it represents focal information, as in the following utterance (Fusellier-Souza, 2004: 285):

(5)

(answering a question about a journey: how many nights/days did it take to go there?)

GO me [pointing at himself] NIGHT [or day: sleeping sign] THREE

### 1.3 Emerging language in the lab

Emerging sign languages and second language acquisition are not the only contexts in which newly emerging language systems can be studied. Recently, an experimental method in which adult participants are asked to convey information using only gesture and no speech has been employed to show that participants in such situations do not rely on the grammar rules of their native language, but use language independent means to structure their gesture sequences (Goldin-Meadow et al., 2008).

Silent gesture experiments have informed the debate about the origins of basic word order (ordering of Subject, Object and Verb) in the languages of the world: a growing number of studies focus on the expression of simple transitive events, typically involving an actor, a patient and an action. Such events, when they involve motion through space, were shown to lead to SOV word order (Goldin-Meadow et al., 2008; replicated in, a.o., Langus & Nespors, 2010; Schouwstra & de Swart, 2014). Given its prominent role in silent gesture, and the fact that it also shows up in emerging sign languages (see Napoli & Sutton-Spence, 2014, and references therein), SOV is taken to be important in the origins of basic word order in human language.

However, SOV is certainly not the only order found in silent gesture. Langus and Nespors (2010) show that for *complex* events such as ‘Man tells boy that girl catches fish’, people use, in silent gesture, an utterance structure typical for SVO languages: they place subordinate clauses after the verb of the main clause. In other words, when the structure of the information in silent gesture becomes more complex, participants do not maintain structures consistent with the SOV ordering that was found for simple events. Another possible reason to move away from SOV ordering is when events are reversible. Two separate publications show that reversible events, transitive events with two animates (such as ‘girl pushes boy’), show a greater preference for SVO order cross-linguistically (Hall et al., 2013; Gibson et al., 2013). Also *intensional* events make people deviate from SOV order (Schouwstra & de Swart, 2014). A dramatic increase in SVO usage was found for these events, in which the patient is either dependent on the verb (‘witch draws banana’) or non-specific (‘pirate looks for ball’) or non-existent (‘princess dreams of vase’), compared to *extensional* events (events like those used by Goldin-Meadow et al., 2008), for which SOV was still preferred.<sup>3</sup>

Schouwstra & de Swart (2014) postulate the possibility that silent gesture gives rise to,

---

<sup>3</sup> The distinction between intensional and extensional verbs is a well defined (although also thoroughly discussed) distinction in formal semantics: the arguments of intensional verbs are defined in terms of their intension (meaning), whereas arguments of extensional verbs are defined in terms of their extension (an object in the world). See Forbes (2013) for a more detailed account.



instead of rigid SOV ordered utterances plus some exceptions, a *flexible* ordering of elements, based on the semantic properties of the information that is described plus the cognitive biases of individuals (such as the preference to present information that is cognitively more basic before more abstract or relational information).

This paper aims to establish a closer connection between natural and lab data, by taking the expression of temporal structure, as observed in the Basic Variety and homesign, to a laboratory environment. What happens when participants describe events that are temporally displaced, using silent gesture?

## **2. Experiment: temporal events in silent gesture**

### **2.1 Introduction**

To test if the temporal utterance structures from homesign and the Basic Variety can be replicated in the lab, naïve participants were asked to use silent gesture to convey information about temporally displaced events. The stimuli were (line drawings of) simple events taking place at a given time, such as ‘a gnome eats pizza at three o’clock’. In Dutch (the native language of all participants) such events can be described in three relevant ways:

(7)

*Om drie uur eet de kabouter pizza.*

At three o'clock eats the gnome pizza.

(TVSO)

(8)

*De kabouter eet om drie uur pizza.*

The gnome eats at three o'clock pizza.

(SVTO)

(9)

*De kabouter eet pizza om drie uur.*

The gnome eats pizza at three o'clock.

(SVOT)

Note that in these sentences, tense information is not so much conveyed by inflection on the verb, but rather by the adverbial. The adverbial, however, can be inserted in different places in the sentence (before or after the propositional information, or even in the middle), and the structure of the sentence is in some cases affected by the presence of the adverbial. In sentence (7), the order of the propositional information is affected by the presence of the adverbial: because Dutch is a V2 language, the sentence-initial adverbial causes the word order of the remainder of the sentence to change into VSO.

If silent gesture makes people by-pass the rules of their native language, these patterns should *not* surface in a silent gesture experiment with temporally displaced events. And if the silent gesture domain triggers strategies similar to those in the Basic Variety and homesign, one would expect to see the propositional and temporal information presented separately, with the temporal information utterance-initial. On the other hand, however, there is the observation (Langus & Nespors, 2010) that complex events lead to structuring different from that of simple events. Temporally displaced events are more complex than simple transitive events, so one could also expect the presence of temporal information to affect the ordering of the elements in an utterance.

## 2.2 Materials, participants, procedure

A set of 16 temporal events was used: line drawings of simple events, with a clock in the upper right corner. An example is, as mentioned above, ‘Gnome eats pizza at three o’clock’; see **Fig. 1**. Seven out of the 16 temporal events were extensional, and 9 were intensional events. Each picture had a clock showing a time between one and ten o’clock. To control the order in which participants looked at the different kinds of information on the picture, this information was presented in stages, in two possible orders: clock-first items presented a picture of a clock first (in the upper right corner), and then the picture of the event was added to obtain the full picture. Event-first items presented the event first, and then the clock was added to obtain the full picture.

16 Dutch participants (9 male, 7 female, age range 17—26) with no knowledge of any existing sign languages took part in the experiment. They received €5 for participating. The participants were first asked to describe the pictures in spoken Dutch, using simple declarative sentences. Their responses were recorded. Subsequently, they were asked to describe the stimuli using only gesture and no speech. The stimuli were presented on a laptop screen, in random order, and the participants’ gestures were recorded. The researcher, or a laboratory assistant, was present during the experiment; they sat across the participant without being able to see the screen. It was pointed out to the participants that the experimenter knew all stimuli, but did not know in which order they would be presented. Practice trials consisted of two temporal events (one intensional, one extensional; one was presented clock-first, one was presented event-first). During the practice trials, the experimenter encouraged the participant to describe as many details as they could (without explicitly referring to any of the elements of the event); during the test phase, the experimenter gave no spoken feedback.



### 2.3 Predictions

If participants indeed apply the temporal displacement strategy (similarly to the Basic Variety and homesign), one would expect participants to be more likely to put the temporal adverb in a frontal position in gesture than in speech. If participants indeed by-pass the grammatical rules of their native language, they should not use the word order TVSO that is specific for Dutch.

To see if the structure of silent gesture changes when participants describe complex events (compared to simple events), the word order for the propositional information is relevant.

Assuming that the pattern found in Schouwstra & de Swart (2014) extends to more complex utterances one should expect a preference for SOV word order for extensional events, and an SVO preference for intensional events.

To check if participants were influenced by the order in which the information (temporal and propositional information) is presented, word orders for clock-first items are compared with

event-first items. If participants are primed by the order of presentation, one should see an effect of item type on gesture order.

## 2.4 Analysis and results

The speech strings from the first part of the experiment were transcribed and coded for word order. The video recordings from the second part (256 gesture strings) were coded for word order. The items for which a participant had started gesturing before all the information (time *and* event) was visible (5 items) were excluded. The remaining 251 items were coded for word order by two independent coders. Multiple consecutive gestures that denoted the same entity were considered as one constituent. For example, a gesture sequence [*tap-wrist* + *two fingers* + *moustache* + *hat* + *triangle* + *eat*] for the example in figure 1 would be coded as TSOV (*tap-wrist* + *two fingers* form the temporal adverbial; *moustache* + *hat* denote the Subject; *triangle* denotes the Object; *eat* denotes the Verb). The two coders reached 94.8% agreement, and all items for which the two coders did not have matching orders (13 of 251 items) were excluded. The remaining 238 items were used for statistical analysis.

### Speech versus gesture

The placement of the temporal information in the speech strings was compared to that in gesture strings. A repeated measures ANOVA shows a significant main effect for modality on the placement of the temporal adverbial  $F(1,15) = 15.940$ ,  $p = .001$ . Pairwise Bonferroni corrected comparisons show that the temporal adverb was placed in the initial position more often when participants were gesturing ( $M = .629$ ,  $SE = .113$ ) than when they were speaking ( $M = .155$ ,  $SE = .086$ ).

The inversion of word order (from SVO to VSO) after an utterance-initial adverbial in speech was compared to that in gesture. In 39 (of 256) speech strings, the temporal adverb was

placed in the initial position, and in all of these, the word order was TVSO. Of the gesture strings, 158 (of 238) had the temporal adverb in initial position, and in only 2 occurrences was the specific order TVSO (the dominant orders were TSOV and TSVO).

There is a significant difference between the two modalities with respect to the occurrence of TVSO order in those strings where T is put in initial position  $\chi^2(1)=179.076$ ,  $p < .0001$ .

#### Intensional and extensional events

It was investigated whether the difference between intensional and extensional events leads to different orderings. A repeated measures ANOVA shows a significant interaction between the gesturing order of the event information and the kind of event (intensional or extensional).

$F(1,14)=37.062$ ,  $p<.0001$ . Pairwise Bonferroni corrected comparisons show that participants use SOV order more often for extensional events ( $M=.581$ ,  $SE=.093$ ) than for intensional events ( $M=.066$ ,  $SE=.015$ ).

#### Influence of order of presentation

It was verified whether the order of presentation of the information (temporal information and event information) in the items had an influence on the placement of the temporal adverbial: did participants change the placement of the temporal adverbial in their gesturing if the information was presented in a different order? Using a repeated measures ANOVA, no significant main effect of order of presentation of the information (temporal information first vs. event information first) in the pictures was found on the placement of the temporal adverbial  $F(1, 15)=1.931$ ,  $p=.185$ .

## 2.5 Discussion

The main result, the fact that participants were more likely to place the temporal adverbial utterance-initially in gesture than in speech, shows that participants engaged in silent gesture

behave differently from when they speak their native language, and similarly to the patterns observed in the Basic Variety and homesign. This picture is strengthened by the observation that the typical verb-second pattern found in spoken Dutch (resulting in VSO order when the adverb is initial) is not observed in silent gesture.<sup>4</sup>

A second way in which this paper shows that in silent gesture the presence of temporal information does not influence the order of the remaining information is by replicating the semantic basic word order patterning, SOV for extensional and SVO intensional events, from Schouwstra & de Swart (2014). This suggests even more strongly that in situations where people cannot use patterns from any conventional language they know, they shape their utterances according to their semantic properties, even when they convey different kinds of information (temporal and propositional in this case).

Finally, it should be noted that many factors may influence word order in silent gesture, such as the communicative setup (Gershkoff-Stowe & Goldin-Meadow, 2002), specific production based constraints (Hall et al., 2013), or the presence of a lexicon (Marno et al., 2015).

However, the word order patterns observed in the present study are strikingly similar to the patterns observed in homesign and Basic Variety, and to their semantic or conceptual structure. Looking at the role of semantic structure in emerging language systems and its interaction with other factors that determine language structure is an interesting direction for future research (Schouwstra, Smith & Kirby, in prep).

### **3. General discussion: expression of temporal information with restricted means**

It is well known that most existing languages have sophisticated means for conveying temporal and aspectual information, and that the tense/aspect systems of people's native language influences the way in which they conceptualise events. For example, it determines

---

<sup>4</sup> It would be interesting to know whether this pattern exists in the Basic Variety of learners of Dutch.

the extent to which they specify aspectual information when they process information (Bott & Hamm, 2014), or the likelihood with which they mention endpoints when they describe events (Carroll et al., 2004). These results suggest that temporal information is highly interwoven with other information in human cognition (Bott & Hamm, 2014). But what happens when one is *not* able to use one's native language, in situations of communicative stress, in which individuals build a communication system from scratch (or from the limited lexical means they have available)? To see any tense/aspect emerge at all in such situations is interesting, because it tells us something about the importance of referring to the non-present. And the particular patterns that emerge consistently in such situations may tell us something about the language independent cognitive biases that form the basis of temporal structure in language.

This paper started with observations from the Basic Variety and homesign, in which speakers typically refer to the non-present by using an utterance-initial temporal adverbial.

Semantically, the function of an utterance-initial adverbial is to create a temporal frame, a place on the timeline in the context of which the rest of the information is interpreted (de Swart, 1999), and the utterance structures found in the BV and homesign mirror that pattern. The experimental study presented in this paper confirms that silent gesture elicits the same strategy: participants present the temporal information utterance-initially, and separately from the propositional information.

When time and tense is studied in existing languages, the emphasis is mostly on verbal inflection, and to a lesser extent on adverbs (Klein, 2008). However, the special role of temporal adverbials has been noticed, not only in the literature about second language acquisition and homesign mentioned in this paper, but also in language processing studies. In an ERP study that investigates sentences with an adverb/tense marking mismatch (such as '*Last Sunday, Vincent paints the window frames of his country house.*'), Baggio (2008) showed that the conflicting semantic information is detected rapidly (i.e. before the syntactic



parsing has completed), and he found suggestive evidence that participants semantically readjust the temporal constraints set up by the verb, by arriving at a ‘narrative present’ reading of the sentence. That would mean that the information coming from the adverb is dominant over the verbal inflection.<sup>5</sup> The results presented in the present paper show that fronting an adverbial is a powerful way to indicate temporal displacement in situations without a conventional language. Even though full languages usually have complex means to convey many subtleties of temporal structure, traces of this more basic strategy may still be present in language processing. It is recognised, however, that more language processing work is needed, for example to investigate the kind of utterances that were central in this paper: with a temporal adverbial, but no inflection on the verb (Gullberg & Indefrey, 2008).

Although the silent gesture paradigm and natural data from restricted systems have been connected before (Goldin-Meadow, 2015; Goldin-Meadow & Brentari, 2015), this is the first successful attempt to replicate in the laboratory a particular kind of structure observed cross-modally in natural data. Silent gesture offers a level of control and precision that is hard to match in collecting natural data, and this makes it possible to ask fine-grained questions about emerging language situations. However, the existing systems like the Basic Variety and homesign, despite their disadvantages, offer something that the lab task cannot offer: natural circumstances. For that reason, it is good to conceive of the silent gesture paradigm as a way to complement rather than replace the evidence from existing restricted linguistic systems. Combining the precision of laboratory experiments with the naturalness of field data is a promising approach in uncovering the cognitive processes that shape emerging language (Motamedi et al., in press).

---

<sup>5</sup> Relatedly, Farوقي-Shah and Dickey (2008) found that in an online sentence judgement task, participants were slower to judge adverb/tense mismatches when the adverb was pre-posed (‘Last year, my sister lived in New Hampshire’) than when the adverb was post-posed (‘My sister lived in New Hampshire last year’). This result was found in their control group, and because the paper focuses mainly on the aphasic target group, the result remains largely unexplained. It could, however, mean that a pre-posed adverbial has a stronger effect on sentence interpretation and is therefore harder to evaluate than a post-posed adverbial.

## Acknowledgements

The author would like to thank Kimberly Snoyl for her help in running the experiment and coding the videos, Marianne Smit and Nicky Marien for their help with the stimuli, and Henriëtte de Swart for thoughtful comments on an earlier version of the paper. The author was funded by the British Academy (grant number pf130034); the experiment was carried out while the author was employed at the Utrecht Institute of Linguistics OTS, Utrecht University.

## References

- Baggio, G. (2008). Processing temporal constraints: An ERP study. *Language Learning*, 58(s1), 35-55.
- Benazzo, S. (2009). The emergence of temporality. In R. Botha & H. de Swart (Eds.), *Language evolution: the view from restricted linguistic systems*. LOT.
- Bott, O. and Hamm, F. (2014). Crosslinguistic variation in the processing of aspect. In B. Hemforth, B. Schmiedtová, and C. Fabricius-Hansen (Eds), *Experimental approaches to crosslinguistic meaning*. Springer, Studies in theoretical psycholinguistics.
- Bybee, J. L., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: The University of Chicago Press.
- Carroll, M., Von Stutterheim, C., & Nuese, R. (2004). The language and thought debate: A psycholinguistic approach. *Trends in linguistic studies and monographs*, 157, 183-218.
- Dahl, Ö. (1985). *Tense and aspect systems*. New York: Basil Blackwell.
- Dahl, Ö., & Velupillai, V. (2013). The past tense. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Faroqi-Shah, Y., & Dickey, M. W. (2009). On-line processing of tense and temporality in agrammatic aphasia. *Brain and Language*, 108(2), 97-111.

Forbes, G. (2013). Intensional Transitive Verbs. Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition).

<http://plato.stanford.edu/archives/fall2013/entries/intensional-trans-verbs>.

Fusellier-Souza, I. (2004). Sémio-genèse des langues des signes. Étude de langues des signes primaires (lsp) pratiquées par des sourds Brésiliens. Unpublished doctoral dissertation, Université Paris VIII, Saint Denis.

Gärdenfors, P. (2004). *How homo became sapiens: On the evolution of thinking*. Oxford University Press.

Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of crosslinguistic word-order variation. *Psychological science*, 0956797612463705.

Goldin-Meadow, S., & Mylander, C. (1998). Spontaneous sign systems created by deaf children in two cultures. *Nature*, 391(6664), 279-281.

Goldin-Meadow, S. (2003). *Resilience of language*. New York: Psychology Press.

Goldin-Meadow, S., So, W. C., O'zyu'rek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *PNAS*, 105(27), 9163–9168.

Goldin-Meadow, S. (2015). The impact of time on predicate forms in the manual modality: Signers, homesigners, and silent gesturers. *Topics in cognitive science*, 7(1), 169-184.

Goldin-Meadow, S., & Brentari, D. (2015). Gesture, sign and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences*, 1-82.

Gullberg, M., & Indefrey, P. (2008). Cognitive and neural prerequisites for time in language: Any answers?. *Language Learning*, 58(s1), 207-216.

Hall, M. L., Mayberry, R. I., & Ferreira, V. S. (2013). Cognitive constraints on constituent order: Evidence from elicited pantomime. *Cognition*, 129(1), 1–17.

Hamm, F., & Bott, O. (2014). Tense and aspect. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2014 ed.).

Hurford, J. R. (2007). *The origins of meaning* (Vol. I). Oxford University Press.

Klein, W. (2001). Elementary forms of linguistic organization. In J. Trabant & S. Ward (Eds.), *New essays on the origin of language* (pp. 81–102). de Gruyter.

Klein, W., & Perdue, C. (1997). The Basic Variety (or: couldn't natural languages be much simpler?). *Second Language Research*, 13, 301–347.

Langus, A., & Nespors, M. (2010). Cognitive systems struggling for word order. *Cognitive Psychology*, 60(4).

Marno, H., Langus, A., Omidbeigi, M., Asaadi, S., Seyed-Allaei, S., & Nespors, M. (2015). A new perspective on word order preferences: the availability of a lexicon triggers the use of SVO word order. *Frontiers in psychology*, 6.

Millikan, R. G. (2004). On reading signs: Some differences between us and the others. In K. Oller & U. Griebel (Eds.), *Evolution of communication systems: A comparative approach* (p. 15-30). Cambridge, MA: MIT Press.

Morford, J. P., & Goldin-Meadow, S. (1997). From here and now to there and then: The development of displaced reference in homesign and English. *Child development*, 68(3), 420-435.

Motamedi, Y., Schouwstra, M., Kirby, S., (in press). An evolutionary approach to sign language emergence: from state to process. Commentary on S. Goldin-Meadow and D. Brentari, *Gesture, sign and language: the coming of age of sign language and gesture studies*. Behavioral and Brain Sciences.

Napoli, D. J., & Sutton-Spence, R. (2014). Order of the major constituents in sign languages: implications for all language. *Frontiers in Psychology*, 5.

Padden, C., Meir, I., Sandler, W., & Aronoff, M. (2010). Against all expectations: Encoding subjects and objects in a new language. Hypothesis A/hypothesis B: Linguistic explorations in honor of David M. Perlmutter,

383–400.

Schouwstra, M., & Swart, H. de. (2014). The semantic origins of word order. *Cognition*, 131(3), 431–436.

Schouwstra, M., Smith, K. & Kirby, S. (in prep). The emergence of word order conventions: improvisation, interaction and transmission. Manuscript in preparation.

Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, 305(5691), 1779–1782.

Starren, M. (2001). *The second time. the acquisition of temporality in Dutch and French as a second language*. Doctoral dissertation, Netherlands Graduate School of Linguistics.

Steedman, M. (1994). Temporality. In *Handbook of logic and language*. Elsevier Science B.V.

Swart, H. de (1999). 18 Position and Meaning: Time Adverbials in Context. *Focus: Linguistic, cognitive, and computational perspectives*, 336.

## Figures

**Figure 1.**

**Caption:** Fig. 1: Example of one of the stimuli: ‘Gnome eats pizza at three o’clock.’

